

BIG DATA HADOOP AND SPARK DEVELOPMENT

ASSIGNMENT – 14

Table of Contents:

1. Introduction	1
2. Objective	1
3. Problem statement	1
4. Expected Output	
• Task 1	4
• Task 2	8

BIG DATA HADOOP AND SPARK DEVELOPMENT

1. Introduction

In this assignment, the given task is performed and Output of the task is performed and Screenshots are attached.

2. Objective

This assignment consolidates the deeper understanding of the Session – 14 Introduction to Scala, which stands for Scalable Language.

3. Problem Statement

- **Task 1**

1. Given a list of strings - List[String] ("alpha", "gamma", "omega", "zeta", "beta") Find count of all strings with length 4.
2. Convert the list of string to a list of integers, where each string is mapped to its corresponding length.
3. Find count of all strings which contain alphabet 'm'. - Find the count of all strings which start with the alphabet 'a'.

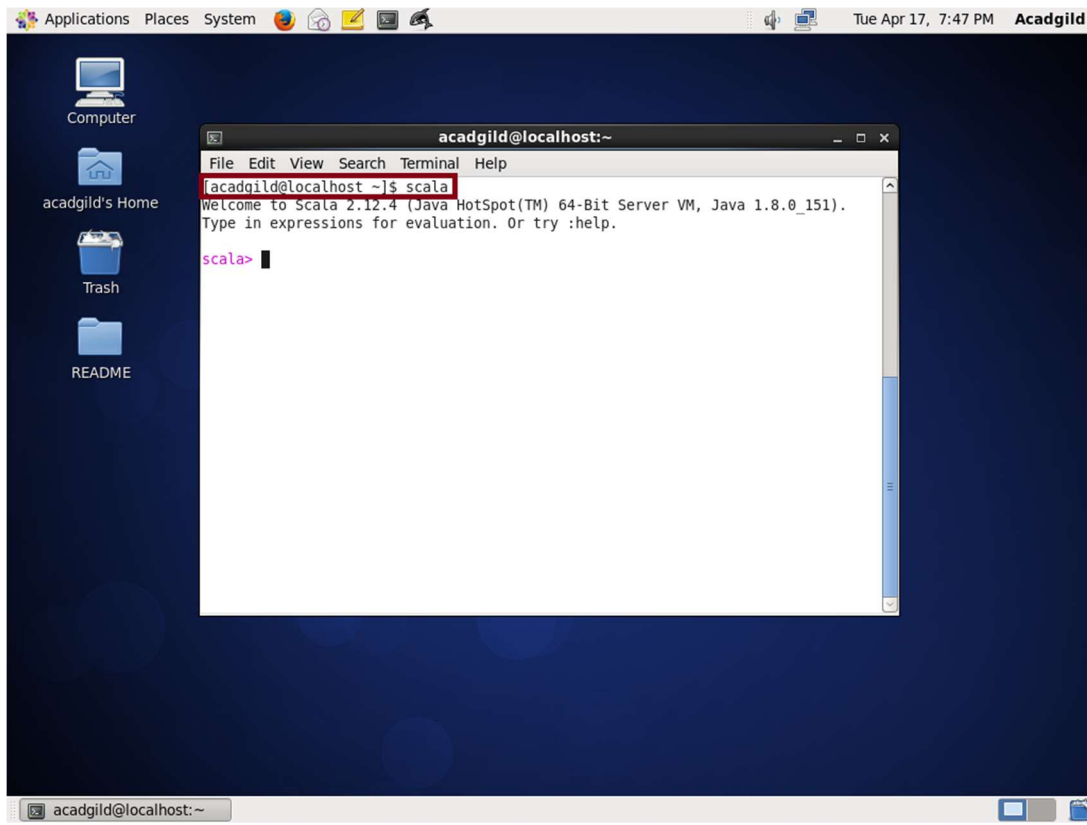
- **Task 2**

1. Create a list of tuples, where the 1st element of the tuple is an int and the second element is a string. Example - ((1, 'alpha'), (2, 'beta'), (3, 'gamma'), (4, 'zeta'), (5, 'omega'))
2. For the above list, print the numbers where the corresponding string length is 4.
3. find the average of all numbers, where the corresponding string contains alphabet 'm' or alphabet 'z'.

4. Expected Output

Preparing to perform tasks:

Start Scalable language console by **scala** command.



The screenshot shows a Linux desktop environment with a dark blue background. On the left side, there are icons for 'Computer', 'acadmild's Home', 'Trash', and 'README'. The top panel displays 'Applications', 'Places', 'System', and the date 'Tue Apr 17, 7:47 PM' next to the username 'Acadmild'. A terminal window titled 'acadmild@localhost:~' is open in the center. The terminal has a menu bar with 'File', 'Edit', 'View', 'Search', 'Terminal', and 'Help'. The command prompt shows the user has entered 'scala' at the shell prompt. The output of the command is: 'Welcome to Scala 2.12.4 (Java HotSpot(TM) 64-Bit Server VM, Java 1.8.0_151). Type in expressions for evaluation. Or try :help.' The prompt now shows 'scala>' with a cursor.

```
acadmild@localhost:~$ scala
Welcome to Scala 2.12.4 (Java HotSpot(TM) 64-Bit Server VM, Java 1.8.0_151).
Type in expressions for evaluation. Or try :help.

scala>
```

❖ Task 1

- Given a list of strings - List[String] ("alpha", "gamma", "omega", "zeta", "beta") - Find count of all strings with length 4.

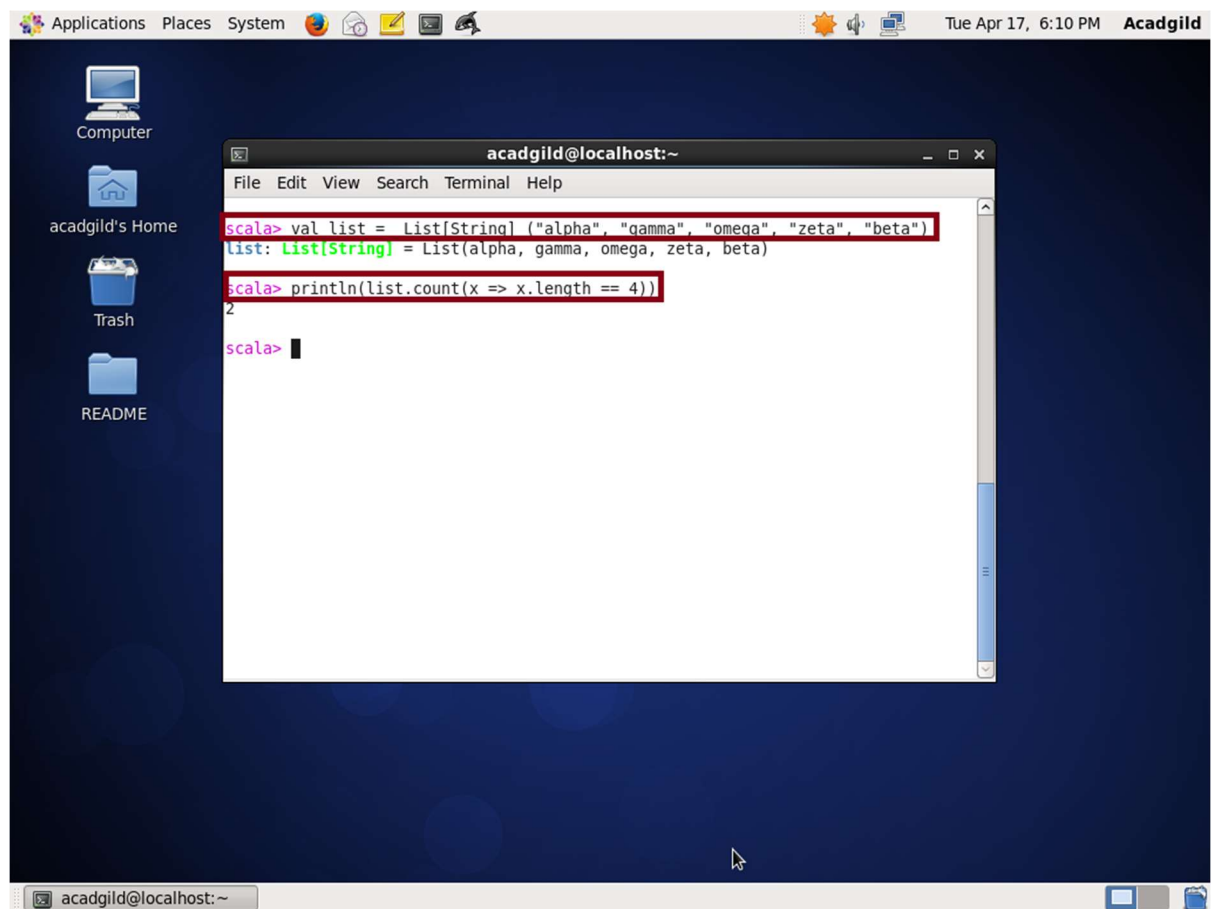
List of string is declared as

```
val list = List[String]("alpha", "gamma", "omega", "zeta", "beta")
```

- To find the count of all string with length 4

By using following command, the count of all strings with length 4 is printed.

```
println(list.count(x=>x.length==4))
```



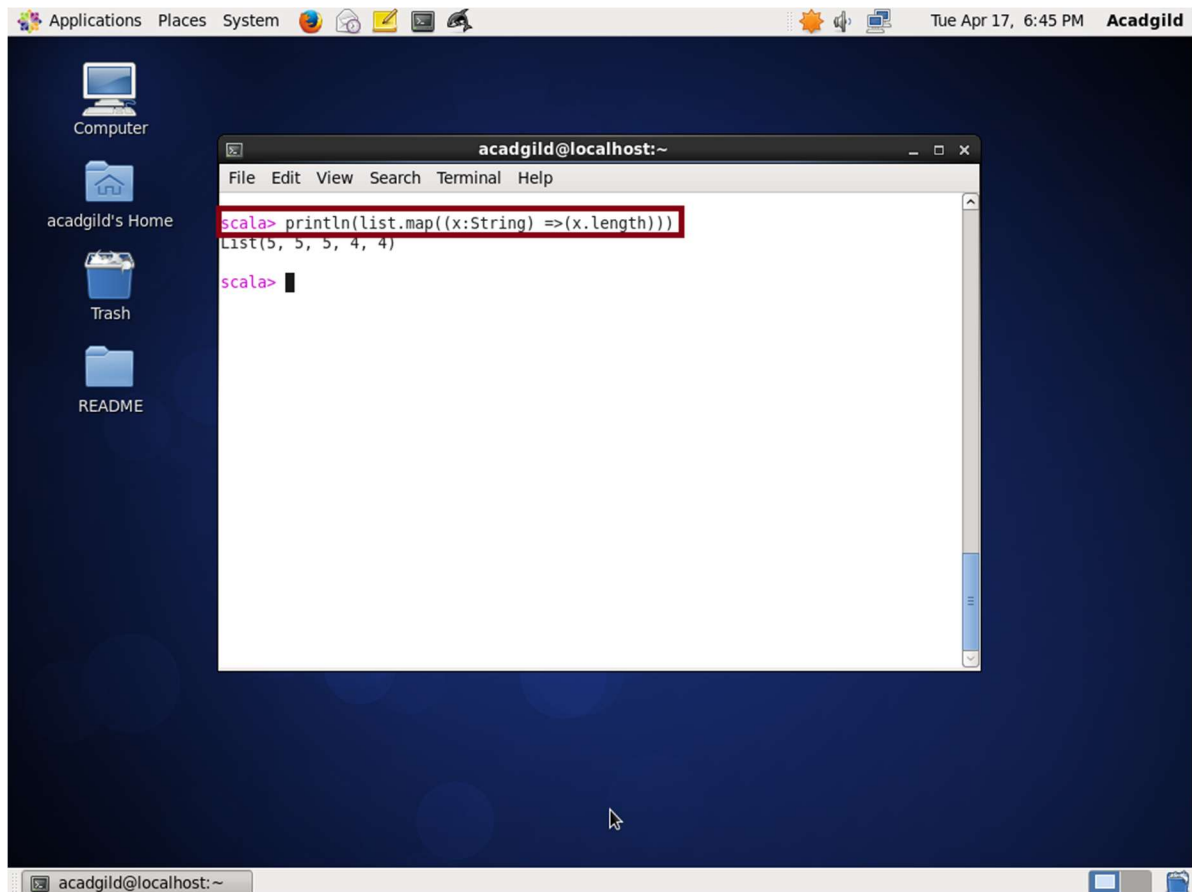
The screenshot shows a Linux desktop environment with a dark blue background. On the left side, there is a sidebar with icons for 'Computer', 'acadgild's Home', 'Trash', and 'README'. The top of the window displays a menu bar with 'Applications', 'Places', and 'System', along with system status icons and the date 'Tue Apr 17, 6:10 PM' and the username 'Acadgild'. A terminal window titled 'acadgild@localhost:~' is open in the center. The terminal has a menu bar with 'File', 'Edit', 'View', 'Search', 'Terminal', and 'Help'. The terminal content shows the following commands and output:

```
scala> val list = List[String] ("alpha", "gamma", "omega", "zeta", "beta")
list: List[String] = List(alpha, gamma, omega, zeta, beta)
scala> println(list.count(x => x.length == 4))
2
scala>
```

- **Convert the list of string to a list of integers, where each string is mapped to its corresponding length.**

The list of string is converted to a list of integers, where each string is mapped to its corresponding length is printed by the following command

```
println(list.map((x:string)=>(x.length)))
```



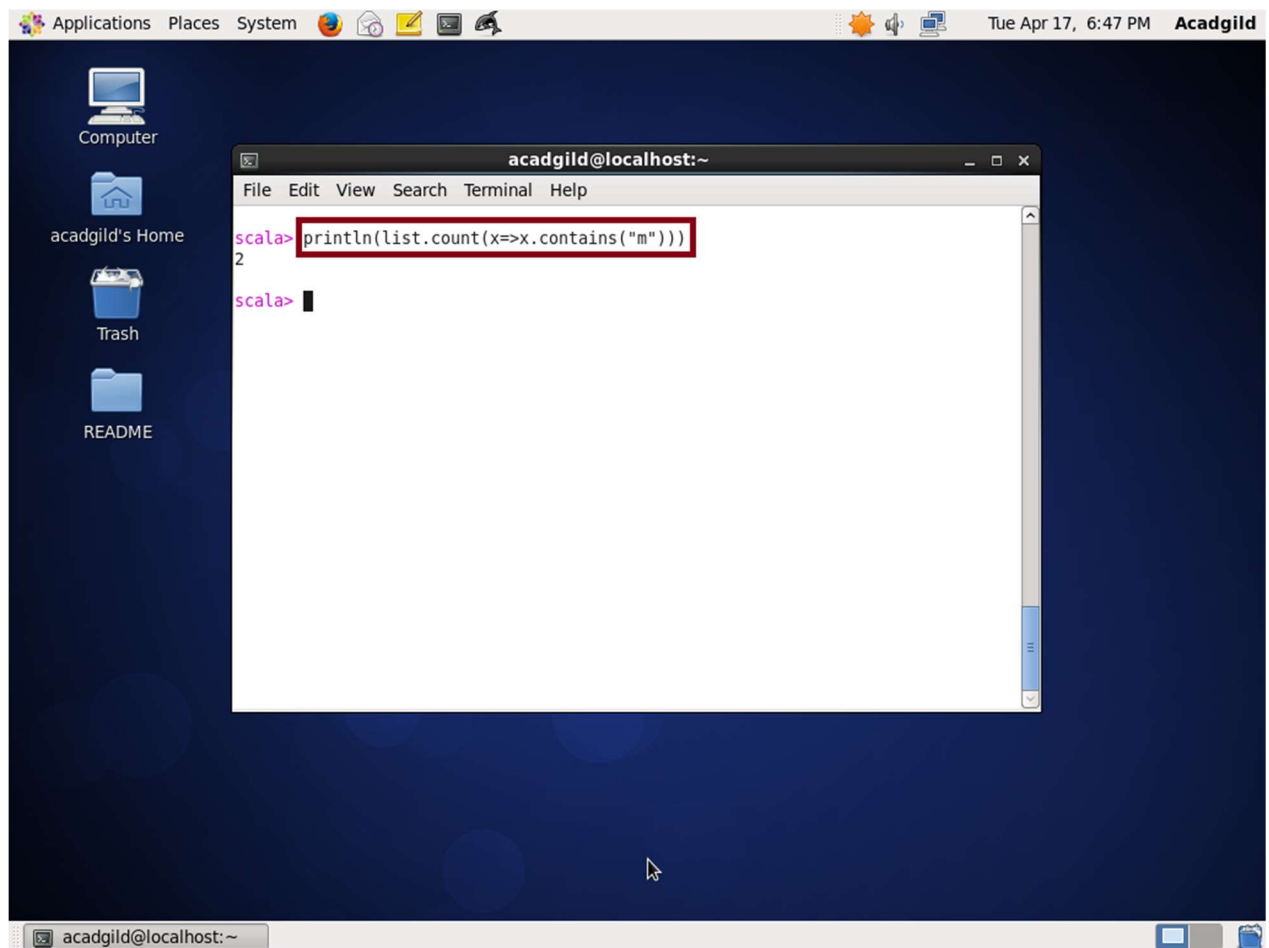
The screenshot shows a Linux desktop environment with a dark blue background. On the left side, there are icons for 'Computer', 'acadgild's Home', 'Trash', and 'README'. At the top, there is a menu bar with 'Applications', 'Places', and 'System'. The system status bar at the top right shows 'Tue Apr 17, 6:45 PM' and 'Acadgild'. A terminal window titled 'acadgild@localhost:~' is open in the center. The terminal has a menu bar with 'File', 'Edit', 'View', 'Search', 'Terminal', and 'Help'. The terminal content shows the following commands and output:

```
scala> println(list.map((x:String) =>(x.length)))
List(5, 5, 5, 4, 4)
scala>
```

- Find count of all strings which contain alphabet 'm'.

By using the following command, the count of strings in the list which contains "m" are printed.

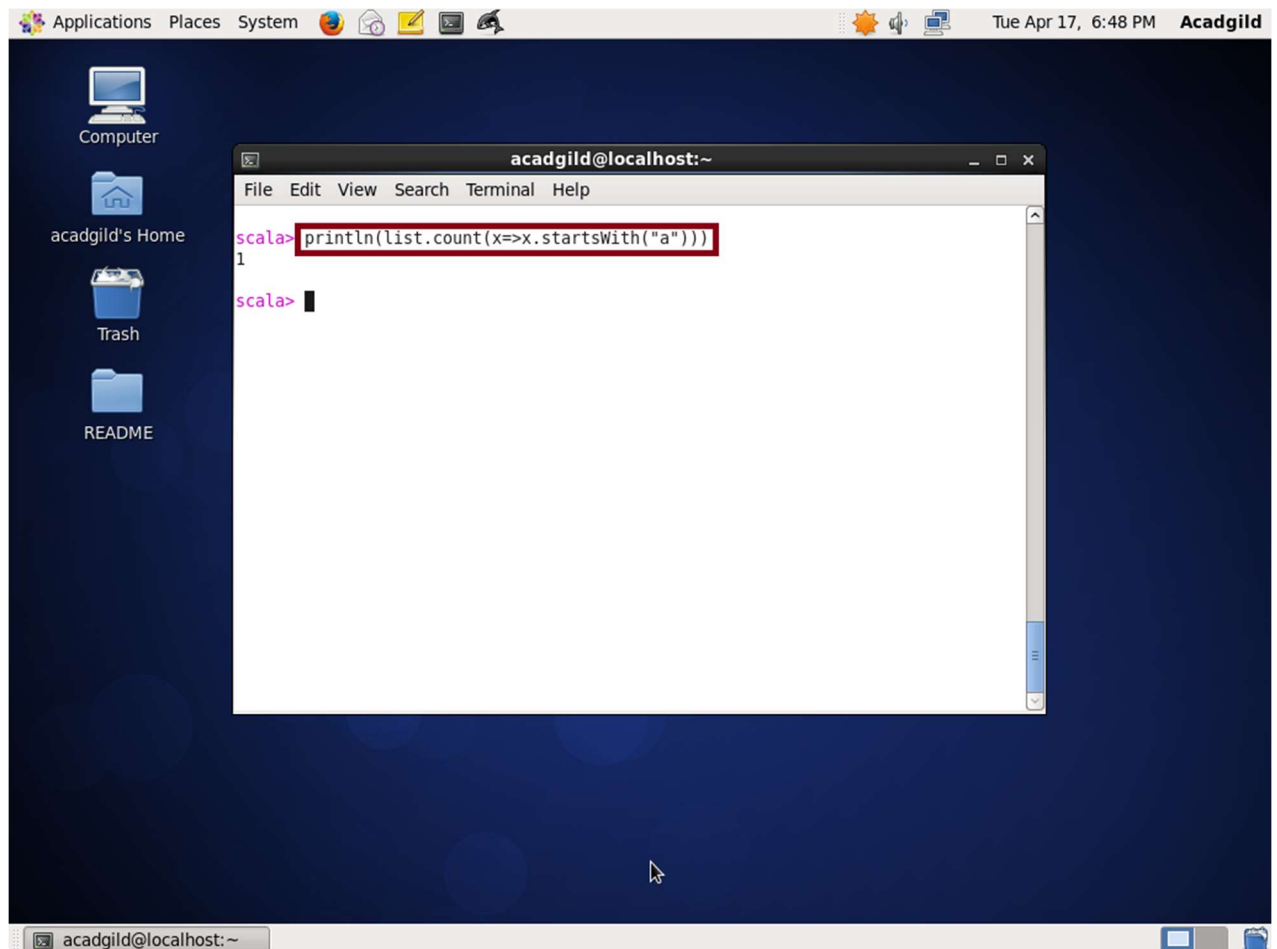
```
println(list.count(x=>x.contains("m")))
```



- Find the count of all the strings which start with the alphabet 'a'.

By using the following command, the strings which starts with the alphabet 'a' are counted and printed.

```
println(list.count(x=>x.startsWith("a")))
```

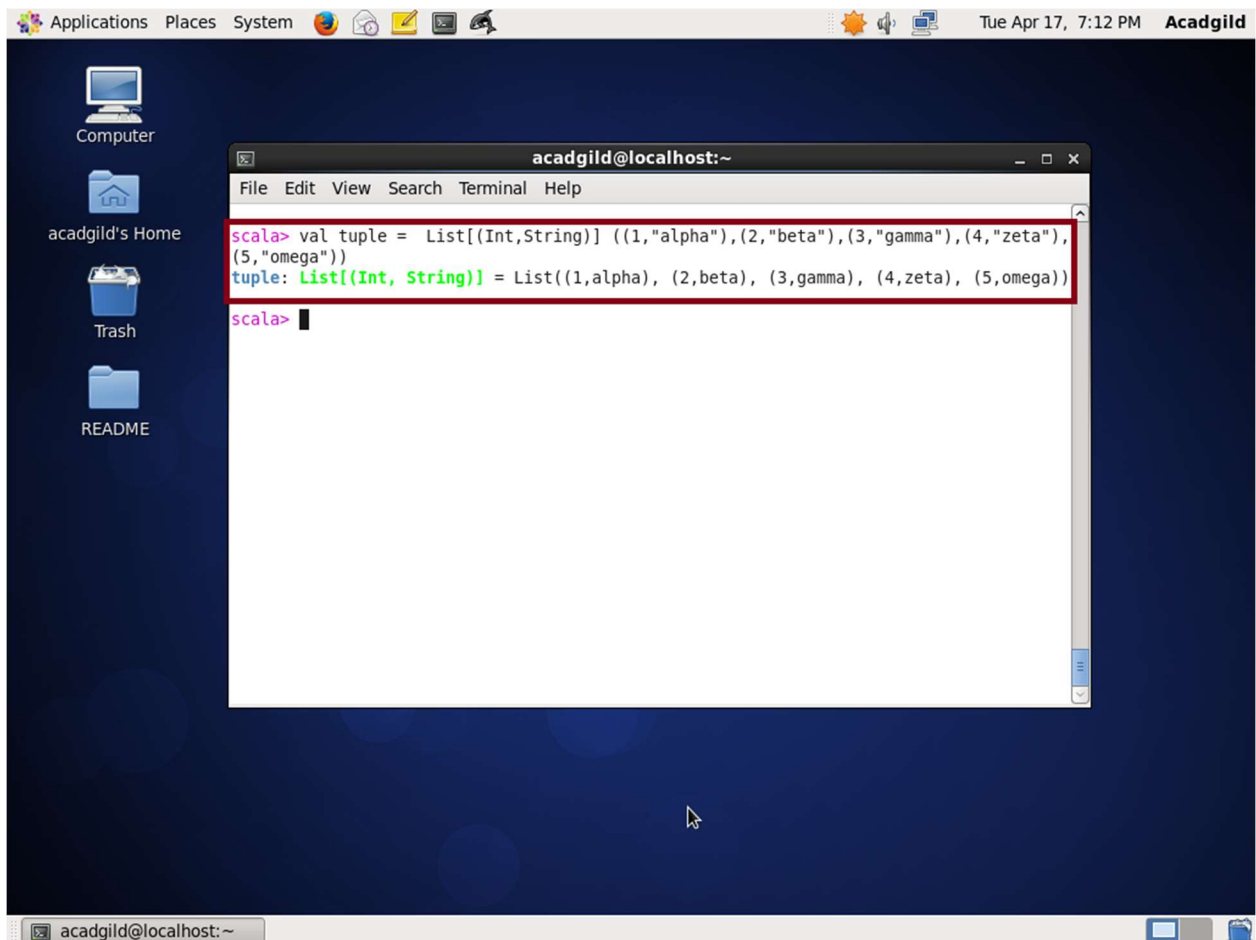


❖ Task 2

- Create a list of tuples, where the 1st element of the tuple is an int and the second element is a string. Example - ((1, 'alpha'), (2, 'beta'), (3, 'gamma'), (4, 'zeta'), (5, 'omega'))

By using the following command, the list of tuples is created, where the 1st element of the tuple is an int and the second element is a string.

```
val tuple=List [(Int, String)] ((1," alpha"), (2," beta"), (3," gamma"), (4," zeta"), (5," omega"))
```



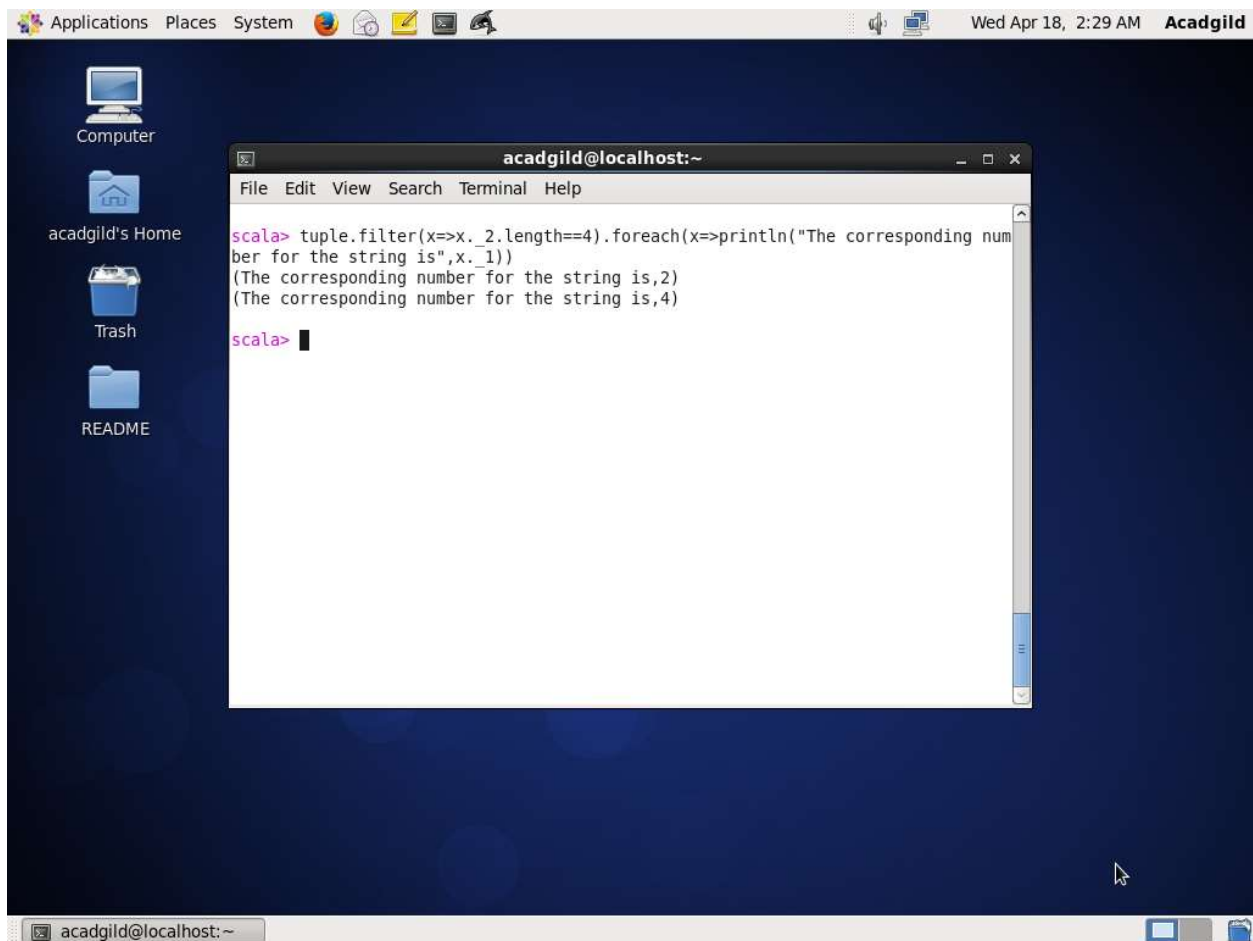
The screenshot shows a Linux desktop environment with a dark blue background. On the left side, there is a sidebar with icons for 'Computer', 'acadgild's Home', 'Trash', and 'README'. The top of the screen features a panel with 'Applications', 'Places', and 'System' menus, along with system status icons and the date 'Tue Apr 17, 7:12 PM' and username 'Acadgild'. A terminal window titled 'acadgild@localhost:~' is open in the center. The terminal has a menu bar with 'File', 'Edit', 'View', 'Search', 'Terminal', and 'Help'. The terminal content shows the following Scala code being entered and executed:

```
scala> val tuple = List[(Int,String)] ((1,"alpha"),(2,"beta"),(3,"gamma"),(4,"zeta"),  
tuple: List[(Int, String)] = List((1,alpha), (2,beta), (3,gamma), (4,zeta), (5,omega))  
scala>
```


- For the above list, print the numbers where the corresponding string length is 4.

By using the following command, the numbers of the corresponding string length is 4 is printed.

```
tuple.filter(x=>(x._2.length==4)).foreach(x=>println("The corresponding number for the string is",x._1))
```

A screenshot of a Linux desktop environment. The desktop background is dark blue. On the left side, there are icons for 'Computer', 'acagdild's Home', 'Trash', and 'README'. At the top, there is a menu bar with 'Applications', 'Places', and 'System'. A terminal window titled 'acagdild@localhost:~' is open in the center. The terminal shows the following command and output:

```
scala> tuple.filter(x=>x._2.length==4).foreach(x=>println("The corresponding number for the string is",x._1))
(The corresponding number for the string is,2)
(The corresponding number for the string is,4)
scala>
```

- Find the average of all numbers, where the corresponding string contains alphabet 'm' or alphabet 'z'.

By using the following command, the strings containing "m" or "z" alphabets are filtered.

```
val tuple_m_z=tuple.filter(x=>(x._2.contains("m")) || (x._2.contains("z")))
```

By using the following command, the length of strings containing "m" or "z" alphabets are found.

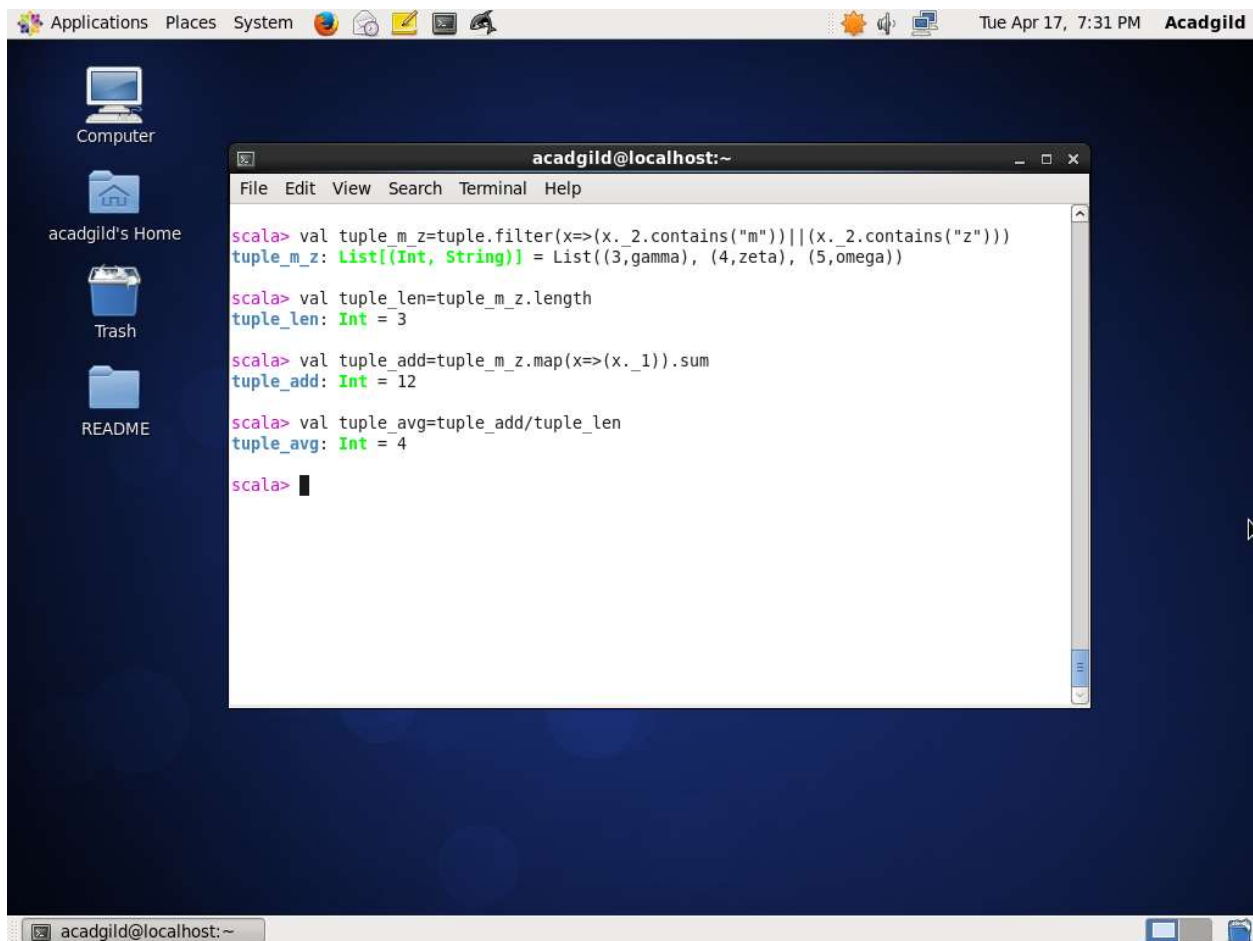
```
val tuple_len=tuple_m_z.length
```

By using the following command, the integers which are mapped to strings containing "m" or "z" alphabets are summed up together.

```
val tuple_add=tuple_m_z.map(x=>(x._1)).sum
```

By using the following command, the average of all numbers, where the corresponding string contains alphabet 'm' or alphabet 'z' are determined.

```
val tuple_avg=tuple_add/tuple_len
```



The screenshot shows a Linux desktop environment with a dark blue background. On the left side, there are icons for 'Computer', 'acadgild's Home', 'Trash', and 'README'. At the top, there is a taskbar with various application icons and the system clock showing 'Tue Apr 17, 7:31 PM' and the username 'Acadgild'. A terminal window titled 'acadgild@localhost: ~' is open in the center, displaying the following Scala code and its output:

```
scala> val tuple_m_z=tuple.filter(x=>(x._2.contains("m")||x._2.contains("z")))
tuple_m_z: List[(Int, String)] = List((3,gamma), (4,zeta), (5,omega))

scala> val tuple_len=tuple_m_z.length
tuple_len: Int = 3

scala> val tuple_add=tuple_m_z.map(x=>(x._1)).sum
tuple_add: Int = 12

scala> val tuple_avg=tuple_add/tuple_len
tuple_avg: Int = 4

scala>
```